REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
16/02/2012	Final	Nov 2008-Jan 2012

4. TITLE AND SUBTITLE

Random matrices, combinatorics, numerical linear algebra and complex networks.

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-09-1-0167

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Van H. Vu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Department of mathematics, Rutgers Univ.
110 Frelinghuysen, Piscataway, NJ 08540.

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Office of Scientific Research,
Mathematics, Information and Life Sciences,
AFOSR/RSL, 875 North RandolphStreet, Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)

AFORS

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Publically Releasable.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Understanding large, random, matrices is important in many areas of interest to AFORS. These includes the probabilistic analysis of problems in numerical linear algebra, the efficiency of the simplex method in linear programming, the key parameters in statistical sampling, the expansion of complex networks such as the Internet graph, to mention a few.

We have developed new methods with combinatorial flavor, combining tools from combinatorics, probability and high dimensional geometry to study several fundamental problems concerning random matrices, motivated by applications on complex networks and data analysis. Computer simulations also play an important role, serving as a guide for theoretical conjectures and as a check for approximations.

15. SUBJECT TERMS

random matrices, complex network, high dimensional data analysis

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Van Vu
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (include area code) 7326891016

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

FINAL REPORT

VAN VU

Grant/Contract Title: RANDOM MATRICES, COMBINATORICS NUMERICAL LINEAR ALGEBRA AND COMPLEX NETWORKS

Grant/Contract Number: FA9550-09-1-0167

1. Condition number: Goldstine-von Neumann problem and Edelman-Spielman-Teng conjecture

Definition 1.1. For an $n \times n$ matrix A, the *condition number* $\kappa(A)$ is defined as

$$\kappa(M) := \|A\|\|A^{-1}\|$$

where $\|\|$ denotes the operator norm. (If A is not invertible, set $\kappa(A) = \infty$.)

Another, and more convenient, way to write $\kappa(M)$ is

$$\kappa(M) = \sigma_1(M)/\sigma_n(M)$$

where $\sigma_1 \geq \cdots \geq \sigma_n(M)$ are the singular values of M. We will use this definition in the rest of the discussion.

The condition number is one of the most important parameters in numerical linear algebra, as it plays a crucial role in running time and accuracy of most linear algebraic algorithms (see, e.g., [1]). As well known, linear algebraic algorithms are of fundamental importance in the study of complex networks, or large data sets in general.

In practice, the smaller the condition number is, the better. Matrices with small condition number (polynomial in n, say) are called well-conditioned while ones with large condition number (super polynomial in n) are called ill-conditioned.

The problem of estimating the condition number of a random matrix was first posed by Goldstine and von Neumann [15, 16], who tried to analyze the performance of one of the very first computers on large matrices in the 1940s.

Let us consider the condition number of a random matrix M_n. For convenience, we assume that the entries of M_n are iid random variables with mean 0 and variance 1. This assumption can be weakened considerably, but notice that it already covers the gaussian model and also the most important discrete model, the Bernoulli (or Rademacher) model (where the entries independently take values ± 1 with probability half).

It is well-known that if we assume furthermore that the entries have light tail (for a more specific statement, see [24]), then the largest singular value σ_1, is strongly concentrated around $2\sqrt{n}$. Since

$$\kappa(M) = \sigma_1(M)/\sigma_n(M),$$

the problem reduces to the study of the least singular value $\sigma_n(M)$.

When the entries of M are standard gaussian, the distribution of σ_n was computed by Edelman [13], confirming the prediction of Goldstine and von Neumann.

Theorem 1.2 (Limiting distributions for gaussian models). *For any fixed $t \geq 0$, we have*

$$\mathbf{P}(n\sigma_n(M_n(\mathfrak{g_R}))^2 \leq t) = \int_0^t \frac{1+\sqrt{x}}{2\sqrt{x}} e^{-(x/2+\sqrt{x})} \, dx + o(1) \tag{1}$$

Edelman's proof used the integration method and relied heavily on concrete formula of the joint distribution of the eigenvalues of random gaussian matrix and thus can not be extended to other cases. On the other hand, both Edelman and Spielman and Teng [27] made an explicit conjecture that the distribution of σ_n is the same in the Bernoulli case.

To illustrate the difficulty of the problem, let us mention that it is already quite non-trivial to show that the least singular value of a random Bernoulli matrix is, with probability tending to one as n tends to infinity, positive. This was done by Komlós in 1967 [19, 20], using combinatorial arguments.

The Pi's first major finding is the confirmation of Edelman-Spielman-Teng conjecture. In fact, he and T. Tao discovered that the limiting distribution of the least singular value (and as a consequence, that of the condition number) of a random matrix is universal, i.e., does not depend on the distribution of the entries.

Theorem 1.3 (Universality for the least singular value). [30] *Let ξ be a random variable with mean 0 and variance 1 and suppose $\mathbf{E}|\xi|^{C_0} < \infty$ for some sufficiently large absolute constant C_0. Let M_n be a random matrix with entries being iid copies of ξ. Then for all $t > 0$, we have*

$$\mathbf{P}(n\sigma_n(M_n)^2 \leq t) = \int_0^t \frac{1+\sqrt{x}}{2\sqrt{x}} e^{-(x/2+\sqrt{x})} \, dx + O(n^{-c}) \tag{2}$$

The implied constants in the $O(.)$ notation depend on $\mathbf{E}|\xi|^{C_0}$ but are uniform in t.

The next figure shows an empirical demonstration of Theorem 1.3 for Bernoulli and for gaussian distributions.
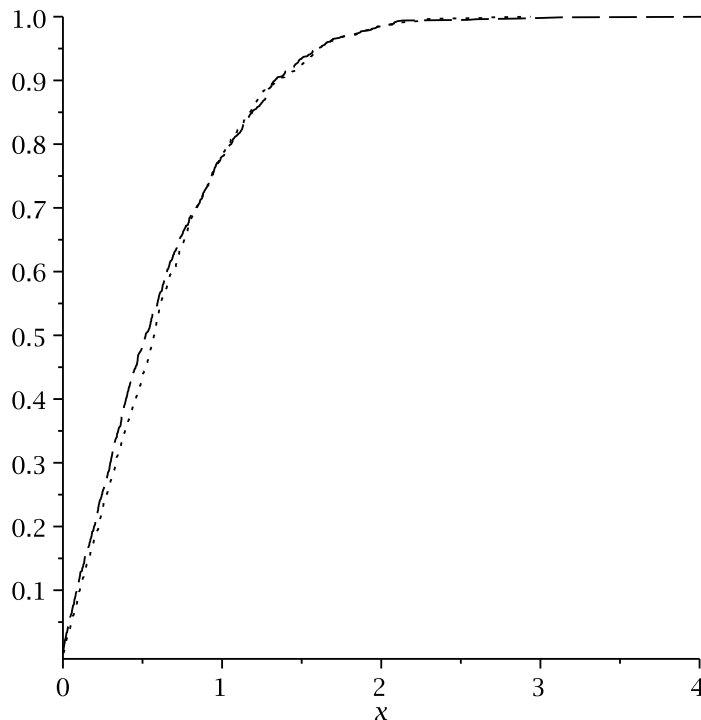
FIGURE 1. Plotted above is the curve $\mathbf{P}(\sqrt{n}\sigma_n(M_n(\xi)) \leq x)$, based on data from 1000 randomly generated matrices with $n = 100$. The dotted curve was generated with ξ a random Bernoulli variable, taking the values $+1$ and -1 each with probability $1/2$; and the dashed curve was generated with ξ a gaussian normal random variable. Note that the two curves are already close together in spite of the relatively coarse data.

Theorem 1.3 also holds in the complex case. This this case the distribution of the least singular value of a (complex) gaussian matrix is also available in Edelman's paper [13].

2. Condition number of randomly perturbed matrices

Now we discuss the general problem of determining the condition number of $A + M_n$, where A can be an arbitrary (deterministic) matrix. This problem is of importance in the study of large data sets with random perturbation. In particular, it played a crucial role in the theory of smoothed analysis introduced by Spielman and Teng [27].

In order to bound the condition number $\kappa(A + M_n)$ from above, we want to bound both $\|A + M_n\|$ and $\|(A + M_n)^{-1}\|$. Bounding $\|A + M_n\|$ is easy. In fact, we expect that in most cases the noise does not change $\|A\|$ significantly. By the triangle inequality

$$\|A\| - \|M_n\| \leq \|A + M_n\| \leq \|A\| + \|M_n\|.$$

On the other hand, in most models of random matrices, one can show that $\|M_n\| = O(\sqrt{n})$ with very high probability, so if $\|A\| = n^{O(1)}$, then $\|A + M_n\|$, with high probability, is polynomially bounded. In practice, high-dimensional matrices usually have not too large entries (compared to their sizes), so this assumption about A is almost always satisfied, and we are going to assume it through our discussion.

The problem is thus to bound the norm of the inverse $\|(A + M_n)^{-1}\|$. Spielman, Sankar and Teng [27, 26] proved the following result for the case when M_n is gaussian (see also [3] for an extension)

Theorem 2.1. *Let A be an arbitrary n by n matrix and M_n be a random matrix with iid gaussian real entries. Then for any $x > 0$,*

$$\mathbf{P}(\|(A + M_n)^{-1}\| \geq x\sqrt{n}) \leq 1.823x.$$

One of the main goal of the proposal was to find an analogue of this result for general random matrices. The problem was posed by Spielman several years ago [27].

The case $A = 0$ was addressed by Rudelson and Vershyin [24], but their proof does not extend to general A. It is not a coincidence, as it has turned out that there is a sharp distinction between the two cases. While in the base case when $A = 0$, the above theorem and [24] showed that the behavior of $\sigma_n(M_n)$ does not depend on the distribution of the entries of M_n (as far as they have mean 0 and variance 1 and bounded moments). On the other hand, when A is not zero, Tao and the PI found an example of A such that the behavior of $\|(A + M_n)^{-1}\|$ when M_n is random Bernoulli *does not* satisfy the bound in Theorem 2.1. This means that one cannot expect Theorem 2.1 holds for non-gaussian variables and more importantly that A should play a role in the inequality.

With this observation in mind, Tao and the PI were able to prove the following theorem [37]

Theorem 2.2. *Let x be a random variable with mean zero and bounded second moment, and let $\gamma 1/2, C0$ be constants. Then there is a constant c depending on x, γ, C such that the following holds. Let M_n be the random matrix of size n whose entries are iid copies of x, A be a deterministic matrix satisfying $\|A\| \leq n^\gamma$ and let $M_n := M + N_n$. Then*

$$\mathbf{P}(s_n(A + M_n) \leq n^{-2(C+1)\gamma}) \leq c(n^{-C+o(1)} + \mathbf{P}(\|N_n\| \geq n^\gamma).$$

This theorem sharpens in extends several existing results, such as those of Tao and the PI [28, 29]) and Rudelson and Vershynin [24].

3. Computing eigenvectors with noise

An important problem that appears in various areas of applied mathematics (in particular statistics, computer science and numerical analysis) is to compute the first few singular vectors of a large matrix. Among others, this problem lies at the heart of PCA (Principal Component Analysis), which has a very wide range of applications (for many examples, see [18, 21] and the references therein).

The basic setting of the problem is as follows:

Problem 3.1. *For a matrix A of size $n \times n$ with singular values $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$, let $v_1, \ldots, v_n$ be the corresponding (unit) singular vectors. Compute $v_1, \ldots, v_k$, for some $k \leq n$.*

Typically n is large and k is relatively small. As a matter of fact, in many applications k is a constant independent of n. For example, to obtain a visualization of a large set of data, one often sets $k = 2$ or 3. The assumption that A is a square matrix is for convenience and our analysis can be carried out with nominal modification for rectangular matrices.

The matrix A, which represents some sort of data, is often perturbed by random noise. Thus, one typically works with $A + M_n$, where M_n, as usual, denotes a random matrix. A natural and important problem is to estimate the influence of noise on the vectors $v_1, \ldots, v_k$. We denote by $v'_1, \ldots, v'_k$ the first k singular vectors of $A + M_n$.

For sake of presentation, we restrict ourselves to the case $k = 1$ (the first singular vector). The following question is of importance

Question 3.2. *When is v'_1 a good approximation of v_1 ?*

A traditional way to measure the distance between two vectors v and v' is to look at $\sin \angle(v, v')$, where $\angle(v, v')$ is the angle between the vectors, taken in $[0, \pi/2]$ (see [14]). (One can also of course use $\|v - v'\|$ but we choose to follow the literature.) To make the problem more quantitative, let us fix a small parameter $\varepsilon > 0$, which represents a desired accuracy. Our question now is to find a sufficient condition for the matrix A which guarantees that $\sin \angle(v_1, v'_1) \leq \epsilon$. It has turned out that the key parameter to look at is the gap (or separation)

$$\delta := \sigma_1 - \sigma_2,$$

between the first and second singular values of A. Classical results in numerical linear algebra yield.

Theorem 3.3. *(Wedin theorem) For any given $\varepsilon > 0$, if $\delta \geq \|E\|/\varepsilon$, then*

$$\sin \angle(v_1, v'_1) \leq \epsilon.$$

In the case when A and $A + M_n$ are Hermitian, this statement is a special case of the famous Davis-Kahan $\sin \theta$ theorem. Wedin [40] extended Davis-Kahan theorem to non-Hermitian matrices, resulting in a general theorem that contains Theorem 3.3 as a special case (see [14, Chapter 8] for more discussion and history).

For most standard models of random matrices, $\|M_n\| \approx 2\sqrt{n}$. So we have

Corollary 3.4. *If $\delta \geq 2\sqrt{n}/\varepsilon$, then with probability $1 - o(1)$*

$$\sin \angle(v_1, v'_1) \leq \epsilon.$$

Low dimensional data and improved bounds.

In a large variety of problems, the data is of small dimension, namely, $r := \text{rank} A \ll n$ (see [12] and the references therein).

The PI and his students run experiments for extremal cases with $r = O(1)$ and see a vast improvement over the bound predicted in corollary 3.4.

.
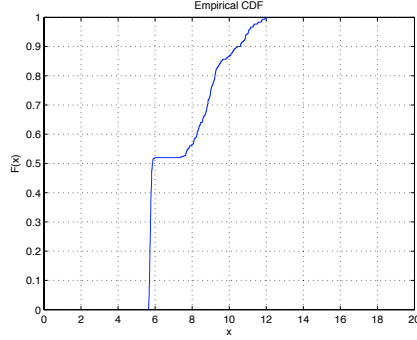
400×400 matrix of rank 2, with gap δ equals 8; this gap is much less than predicted by Wedin's bound. Indeed, we see that with probability close to 1, one can have $\varepsilon \approx 12^o = \frac{1}{15}\pi$. By Corollary 3.4, in order to obtain this accuracy, we need δ to be as large as $2\sqrt{400} \times 15 = 600$.

.
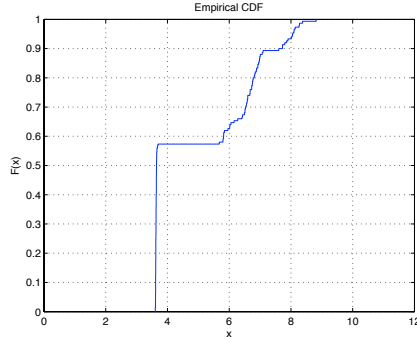
1000×1000 matrix of rank 2, with gap equals 10.

The PI discovered that the bound on δ depends on the *real dimension* r of the data, rather than the size n. He was able to prove [39]

Theorem 3.5. *For any positive constant ϵ there is a positive constant $C = C(\epsilon)$ such that the following holds. Assume that A has rank $r \leq n^{.99}$ and $\frac{n}{\sqrt{r \log n}} \leq \sigma_1$ and $\delta \geq C\sqrt{r \log n}$. Then with probability $1 - o(1)$*

$$\sin \angle(v_1, v_1') \leq \epsilon. \tag{3}$$

Given a desired accuracy ε (say, $1/100$), this theorem means that when r is small, v_1' approximates v_1 under much weaker assumption on the data matrix A, compared to Corollary 3.4.

4. SPECTRUM OF RANDOM GRAPHS

A popular way to model large complex networks is to use random graphs. Naturally, the problem of understanding the spectrum of these graphs is of importance.

The most studied models of random graphs are Erdös-Rényi graph and random regular graph (for definition and more information, see [2]). While these graphs are not yet a good approximation for real-life complex networks,it has turned out that methods developed to study them are very useful in the investigations of more practical models.

Let A_n be the adjacency matrix of the Erdös -Rényi random graph $G(n,p)$; A_n is a random symmetric $n \times n$ matrix whose upper triangular entries are independent identical distributed (iid) copies of a real random variable ξ and diagonal entries are 0; ξ is a Bernoulli random variable that takes values 1 with probability p and 0 with probability $1 - p$.

$$\mathbb{E}\xi = p, \mathbb{V}ar\xi = p(1-p) = \sigma^2.$$

In 1950s, Wigner [41] discovered the famous semi-circle law for the limiting distribution of the eigenvalues of random matrices (see next section). His proof extends, without difficulty, to the adjacency matrix of $G(n,p)$, given that $np \to \infty$ with n.

Theorem 4.1. *For $p = \omega(\frac{1}{n})$, the empirical spectral distribution (ESD) of the matrix $\frac{1}{\sqrt{n}\sigma} A_n$ converges in distribution to the semicircle distribution which has a density $\rho_{sc}(x)$ with support on $[-2, 2]$,*

$$\rho_{sc}(x) := \frac{1}{2\pi} \sqrt{4 - x^2}.$$

If $np = O(1)$, the semicircle law no longer holds. In this case, the graph almost surely has $\Theta(n)$ isolated vertices, so in the limiting distribution, the point 0 has positive constant mass.

Let us now discuss random d-regular graph (denoted by $G_{n,d}$). The spectrum of $G_{n,d}$, was considered by McKay [22] about 30 years ago. He proved that if d is fixed, and $n \to \infty$, then the limiting density function is

$$f_d(x) = \begin{cases} \frac{d\sqrt{4(d-1)-x^2}}{2\pi(d^2-x^2)}, & \text{if } |x| \leq 2\sqrt{d-1}; \\ \\ 0 & \text{otherwise.} \end{cases}$$

This is usually referred to as the McKay or Kesten-McKay law.
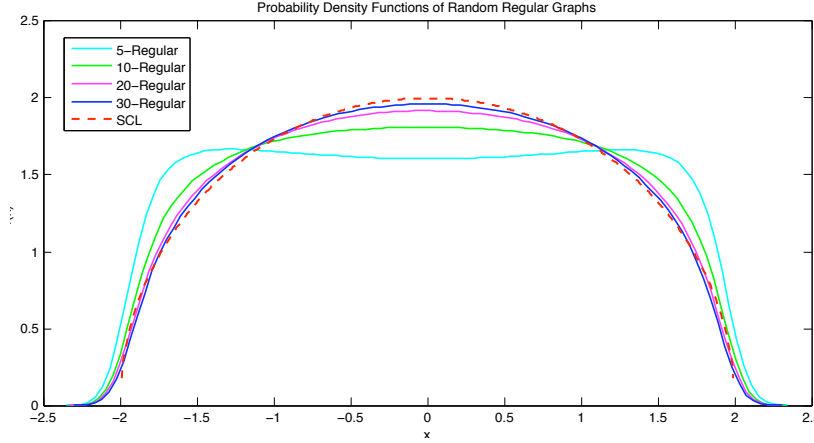
It is easy to verify that as $d \to \infty$, if we normalize the variable x by $\sqrt{d-1}$, then the above density converges to the semicircle distribution on $[-2, 2]$. In fact, a numerical simulation shows the convergence is quite fast (see figure above)

It has been conjectured that Theorem 4.1 holds for $G_{n,d}$ with $d \to \infty$. Let A'_n be the adjacency matrix of $G_{n,d}$, and set

$$M'_n = \frac{1}{\sqrt{\frac{d}{n}(1-\frac{d}{n})}}(A'_n - \frac{d}{n}J).$$

Conjecture 4.2. *If $d \to \infty$ then the ESD of $\frac{1}{\sqrt{n}}M'_n$ converges to the standard semicircle distribution.*

Nothing has been proved about this conjecture, until recently. In [11], Dumitriu and Pal showed that the conjecture holds for d tending to infinity slowly, $d = n^{o(1)}$. Their method does not extend to larger d.

With students L. Tran and L. Wu, the PI were able to established Conjecture 4.2 in full generality. Our method was very different from that of [11] and could be useful in the study of more sophisticated model of random graphs, since it nicely handles a major difficulty that the the edges in the graph are not totally independent (unlike the Erdös-Rényi case).

Theorem 4.3. *If d tends to infinity with n, then the empirical spectral distribution of $\frac{1}{\sqrt{n}}M'_n$ converges in distribution to the semicircle distribution.*

5. LOCAL STATISTICS OF EIGENVALUES

One of the main goal of random matrix theory is to establish limiting distributions concerning the local eigenvalue statistics. To simplify the presentation, we are going to focus on *Wigner Hermitian matrices*, which are perhaps the most prominent model in the field.

Definition 5.1 (Wigner matrices)**.** Let n be a large number. A *Wigner Hermitian matrix* (of size n) is defined as a random Hermitian $n \times n$ matrix M_n with upper triangular complex entries $\zeta_{ij} := \xi_{ij} + \sqrt{-1}\tau_{ij}$ $(1 \leq i < j \leq n)$ and diagonal real entries ξ_{ii} $(1 \leq i \leq n)$ where

- For $1 \leq i < j \leq n$, ξ_{ij}, τ_{ij} are iid copies of a real random variable ξ with mean zero and variance $1/2$.
- For $1 \leq i \leq n$, ξ_{ii} are iid copies of a real random variable $\tilde{\xi}$ with mean zero and variance 1.
- $\xi, \tilde{\xi}$ have exponential decay, i.e., there are constants C, C' such that $\mathbf{P}(|\xi| \geq t^C) \leq \exp(-t), \mathbf{P}(|\tilde{\xi}| \geq t^C) \leq \exp(-t)$, for all $t \geq C'$.

We refer to $\xi, \tilde{\xi}$ as the *atom distributions* of M_n, and ξ_{ij}, τ_{ij} as the *atom variables*. We refer to the matrix $W_n := \frac{1}{\sqrt{n}} M_n$ as the *coarse-scale normalized Wigner Hermitian matrix*, and $A_n := \sqrt{n} M_n$ as the *fine-scale normalized Wigner Hermitian matrix*.

Example. (GUE) An important special case of a Wigner Hermitian matrix is the *gaussian unitary ensemble* (GUE), in which $\xi, \tilde{\xi}$ are gaussian random variables with mean zero and variance $1/2$, 1 respectively. The coarse-scale normalization W_n is convenient for placing all the eigenvalues in a bounded interval, while the fine-scale normalization A_n is convenient for keeping the spacing between adjacent eigenvalues to be roughly of unit size.

Given an $n \times n$ Hermitian matrix A, we denote its n eigenvalues as

$$\lambda_1(A) \leq \ldots \leq \lambda_n(A),$$

and write $\lambda(A) := (\lambda_1(A), \ldots, \lambda_n(A))$.

The study of the eigenvalues $\lambda_i(W_n)$ of (normalized) Wigner Hermitian matrices has been one of the major topics of study in random matrix theory. The properties of these eigenvalues are not only interesting in their own right, but also have been playing essential roles in many other areas of mathematics, such as mathematical physics, probability, combinatorics, and the theory of computing.

A cornerstone of this theory is the *Wigner semicircular law*. Denote by ρ_{sc} the semi-circle density function with support on $[-2, 2]$,

$$\rho_{sc}(x) := \begin{cases} \frac{1}{2\pi}\sqrt{4 - x^2}, & |x| \leq 2 \\ 0, & |x| > 2. \end{cases} \tag{4}$$

Theorem 5.2 (Semi-circular law). *Let M_n be a Wigner Hermitian matrix. Then for any real number x,*

$$\lim_{n \to \infty} \frac{1}{n} |\{1 \leq i \leq n : \lambda_i(W_n) \leq x\}| = \int_{-2}^{x} \rho_{sc}(y) \, dy$$

in the sense of probability (and also in the almost sure sense, if the M_n are all minors of the same infinite Wigner Hermitian matrix), where we use $|I|$ to denote the cardinality of a finite set I.

Remark 5.3. Wigner[41] proved this theorem for special ensembles. The general version above is due to Pastur [25]. The semi-circular law in fact holds under substantially more general hypotheses than those given in Definition 5.1, but we will not discuss this matter further here. One consequence of Theorem 5.2 is that we expect most of the eigenvalues of W_n to lie in the interval $(-2 + \varepsilon, 2 + \varepsilon)$ for $\varepsilon > 0$ small; we shall thus informally refer to this region as the *bulk* of the spectrum.

Theorem 5.2 addressed the global behavior of the eigenvalues. The local properties are much harder and their studies require much more sophisticated tools. Most of the precise theorems have been obtained for the matrices with gaussian entries (GUE).

Notice that in Theorem 5.2, we only need to know the first two moments of the entries (the mean and the variance) to determine the global distribution. Rather surprisingly, Tao and the PI [33] discovered that the local distributions depend only on the first four moments. This led to what we call the Four Moment Theorem.

Definition 5.4 (Matching moments). Let $k \geq 1$. Two complex random variables ξ, ξ' are said to *match to order* k if one has $\mathbf{E}\Re(\xi)^a \Im(\xi)^b = \mathbf{E}\Re(\xi')^a \Im(\xi')^b$ whenever $a, b \geq 0$ are integers such that $a + b \leq k$.

In the model case when the real and imaginary parts of ξ or of ξ' are independent, the matching moment condition simplifies to the assertion that $\mathbf{E}\Re(\xi)^a = \mathbf{E}\Re(\xi')^a$ and $\mathbf{E}\Im(\xi)^b = \mathbf{E}\Im(\xi')^b$ for all $0 \leq a, b \leq k$.

Theorem 5.5 (Four Moment Theorem for eigenvalues). *Let $c_0 > 0$ be a sufficiently small constant. Let $M_n = (\xi_{ij})_{1 \leq i,j \leq n}$ and $M'_n = (\xi'_{ij})_{1 \leq i,j \leq n}$ be two Wigner matrices. Assume furthermore that for any $1 \leq i < j \leq n$, ξ_{ij} and ξ'_{ij} match to order 4 and for any $1 \leq i \leq n$, ξ_{ii} and ξ'_{ii} match to order 2. Set $A_n := \sqrt{n} M_n$ and $A'_n := \sqrt{n} M'_n$, let $1 \leq k \leq n^{c_0}$ be an integer, and let $G : \mathbf{R}^k \to \mathbf{R}$ be a smooth function obeying the derivative bounds*

$$|\nabla^j G(x)| \leq n^{c_0} \tag{5}$$

for all $0 \leq j \leq 5$ and $x \in \mathbf{R}^k$. Then for any $1 \leq i_1 < i_2 \cdots < i_k \leq n$, and for n sufficiently large we have

$$|\mathbf{E}(G(\lambda_{i_1}(A_n), \ldots, \lambda_{i_k}(A_n))) - \mathbf{E}(G(\lambda_{i_1}(A'_n), \ldots, \lambda_{i_k}(A'_n)))| \leq n^{-c_0}. \tag{6}$$

One can use The Four Moment Theorem, combined with various recent results (in particular those of Erdös et. al.; see [7]), to attack several long standing problems. For more discussions, we refer to [7, 17, 33, 35, 36]).

References

[1] D. Bau and L. Trefethen, Numerical linear algebra. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[2] B. Bollobás, Random graphs. Second edition. Cambridge Studies in Advanced Mathematics, 73. Cambridge University Press, Cambridge, 2001.

[3] P. Burgisser, F. Cucker, M. Lotz, The probability that a slightly perturbed numberical analysis problem is difficult, *Math of Computation,* 77, 1559-1583, 2008.

[4] P. Burgisser, F. Cucker, M. Lotz, General formulas for the smooth analysis of condition numbers, *C. R. Acad. Sc. Paris*, 343, 145-150, 2006.

[5] P. Burgisser, F. Cucker, M. Lotz, Smooth analysis of connic condition numbers, *J. Math. Pure et Appl.*, 86, 293-309, 2006.

[6] L. Tran, V. Vu, K. Wang, Sparse random graphs: Eigenvalues and Eigenvectors, *to appear in Random Structures and Algorithms.*

[7] L. Erd**H**os, Universality of Wigner random matrices: a Survey of Recent Results, *arXiv:1004.0861.*

[8] J. Demmel, The probability that a numberical analysis problem is difficult, *Math. Comp.*, 50, 449-480, 1988.

[9] J. Dunagan, D. A. Spielman and S. H. Teng, Smoothed Analysis of the Renegar's Condition Number for Linear Programming, *preprint*.

[10] P. Deif, Orthogonal polynomials and Random matrices: A Rieman-Hilbert approach, AMS 200.

[11] I. Dumitriu and S. Pal. Sparse regular random graphs: spectral density and eigenvectors. Arxiv preprint arXiv:0910.5306, 2009.

[12] C. Eckart and G. Young, The approximation of one matrix by the other of low rank, *Psycometrica*, 1, 211-218, 1936.

[13] A. Edelman, Eigenvalues and condition numbers of random matrices. *SIAM J. Matrix Anal. Appl.* 9 (1988), no. 4, 543–560.

[14] G. Golub and C. Van Loan, Matrix Computations, John Hopkins Univ. Press, 1989.

[15] J. von Neumann (1963), Col lected Works, Vol. V: Design of Computers, Theory of Automata and Numerical Analysis (A. H. Taub, ed.), Pergamon Press, Macmillan, New York.

[16] J. von Neumann and H. H. Goldstine (1947), Numerical inverting of matrices of high order, Bul l. Amer. Math. Soc. 53, 10211099.

[17] A. Guionnet, Grandes matrices aléatoires et théorèmes d'universalité, *Séminaire BOURBAKI*. Avril 2010. 62ème année, 2009-2010, no 1019.

[18] R. Kannan and S. Vempala, Spectral algorithms, Found. Trends Theor. Comput. Sci. 4 (2008), no. 3-4, front matter, 157–288 (2009).

[19] J. Komlós, *On the determinant of* $(0,1)$ *matrices*, Studia Sci. Math. Hungar. **2** (1967) 7–22.

[20] J. Komlós, *On the determinant of random matrices*, Studia Sci. Math. Hungar. **3** (1968) 387–399.

[21] B. Le Roux and R. Henry, Geometric Data Analysis, From Correspondence Analysis to Structured Data Analysis (2004) Dordrecht, Kluwer.

[22] B.D. McKay. The expected eigenvalue distribution of a large regular graph. Linear Algebra and its Applications, 40:203216, 1981.

[23] M.L. Mehta, Random Matrices and the Statistical Theory of Energy Levels, Academic Press, New York, NY, 1967.

[24] M. Rudelson and R. Vershynin, The Littlewood-Offord Problem and invertibility of random matrices, *Advances in Mathematics* 218 (2008), 600–633.

[25] L. Pastur, On the spectrum of random matrices, *Teoret. Mat.Fiz.* **10**, 102-112 (1973).

[26] A. Sankar, S. H. Teng, and D. A. Spielman, Smoothed Analysis of the Condition Numbers and Growth Factors of Matrices, *preprint*.

[27] D. A. Spielman and S. H. Teng, Smoothed analysis of algorithms, *Proceedings of the International Congress of Mathematicians*, Vol. I (Beijing, 2002), 597–606, Higher Ed. Press, Beijing, 2002.

[28] T. Tao and V. Vu, Inverse Littlewood-Offord theorems and the condition number of random discrete matrices, *Annals of Mathematics, to appear*.

[29] T. Tao and V. Vu, The condition number of a randomly perturbed matrix, STOC'07, Proceedings of the 39th Annual ACM Symposium on Theory of Computing, 248–255, ACM, New York, 2007.

[30] T. Tao and V. Vu, Random matrices: the distribution of the smallest singular values, *Geom. Funct. Anal.* 20 (2010), no. 1, 260–297.

[31] T. Tao and V. Vu, Random matrices: Localization of the eigenvalues and the necessity of four moments, *to appear in Acta Math. Vietnamica*.

[32] T. Tao and V. Vu, Random matrices: Universal properties of eigenvectors, to appear in Random Matrices: Theory and Applications

[33] T. Tao and V. Vu, Random matrices: Universality of the local eigenvalue statistics, *Acta Mathematica* 206 (2011), 127-204.

[34] T. Tao, V. Vu, Random covariance matrices: university of local statistics of eigenvalues, *to appear in Annals of Probability*.

[35] T. Tao and V. Vu, The Wigner-Dyson-Mehta bulk universality conjecture for Wigner matrices, *Electronic Journal of Probability*, vol 16 (2011), 2104-2121.

[36] T. Tao and V. Vu, Random matrices: The Universality phenomenon for Wigner ensembles, arXiv:1202.0068.

[37] T. Tao and V. Vu, Smooth analysis of the condition number and the least singular value, *Math. Comp.* 79 (2010), no. 272, 23332352.

[38] L. Tran, V. Vu, K. Wang, Sparse random graphs: Eigenvalues and Eigenvectors, *to appear in Random Structures and Algorithms*.

[39] V. Vu, Singular vectors under random perturbation Random structures and Algorithms, Volume 39, Issue 4, December 2011, Pages: 526538

[40] P. Wedin, Perturbation bounds in connection with singular value decomposition, *Nordisk Tidskr. Informationsbehandling* (BIT) 12 (1972), 99–111.

[41] E.P. Wigner. On the distribution of the roots of certain symmetric matrices. Annals of Mathematics, 67(2):325327, 1958.

DEPARTMENT OF MATHEMATICS, YALE, NEW HAVEN, CT 06520

E-mail address: `van.vu@yale.edu`